2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Evaluation of Mental Workload in Working Memory Tasks with Different Information Types Based on EEG

Kai Guan, Xiaoke Chai, Zhimin Zhang, Qian Li, Haijun Niu*

Abstract—**To explore the effectiveness of using Electroencephalogram (EEG) spectral power and multiscale sample entropy for accessing mental workload in different tasks, working memory tasks with different information types (verbal, object and spatial) and various mental loads were designed based on the N-Back paradigm. Subjective scores, accuracy and response time were used to verify the rationality of the tasks. EEGs from 18 normal adults were acquired when tasks were being performed, an independent component analysis (ICA) based artifact removal method were applied to get clean data. Linear (relative power in Theta and Alpha band, etc.) and nonlinear (multiscale sample entropy) features of EEGs were then extracted. Indices that can effectively reflect mental workload levels were selected by using multivariate analysis of variance statistical approach. Results showed that with the increment of task load, power of frontal Theta, Theta/Alpha ratio and sample entropies at scale more than 10 in parietal regions increased significantly first and decreased slightly then, while the power of central-parietal Alpha decreased significantly first and increased slightly then. Considering the difference between task types, no difference in power of frontal Theta, central-parietal Alpha and sample entropies at scales more than 10 of parietal regions were found between verbal and object tasks, as well as between two spatial tasks. No difference of frontal Theta/Alpha ratio was found in all the four tasks. The results can provide evidence for the mental workload evaluation in tasks with different information types.**

I. INTRODUCTION

The accurate evaluation of mental workload plays a significant role in ensuring the correct exclusion of tasks and the safety of manipulators in human-machine system [1,2]. Frequently used methods of evaluating mental workload include subjective measurement (e.g. NASA-TLX, SWAT, etc.), task performance measurement (e.g. response time, accuracy, etc.) and physiological measurement (e.g. EEG, ECG, etc.) [3]. Compared with subjective and task performance measurements, physiological measurement can objectively, continuously access operator's mental workload without influencing the exclusion of task (Lots of researchers used EEG, ECG, eye movement and other physiological indices to access human's mental workload [4-6]). However, owing to the diversity of the information types in interactive tasks, the issue that different physiological indices are sensitive to different kinds of task is still an obstacle when accessing mental workload in tasks with different information types [7]. Indices that can be effectively used in different tasks have not been proposed.

It was commonly accepted that EEG is the most sensitive physiological signal for its more direct reflection of the process of human's brain compared with ECG, heart rate, eye movement, skin temperature etc. [8]. Previous studies introduced linear features of EEG to access mental workload, including statistic features in time domain [9], spectral power of specific bands in frequency domain [10-11] and so on, among which the power of Theta and Alpha bands were the most commonly used. The conclusions of different researchers, however, didn't reach consensus. Gevins [10] found that the power of frontal Theta increased and parietal Alpha decreased along with the increment of the task load in a verbal and spatial N-Back task; significant difference in Alpha (more specifically Alpha II: 10.5-13Hz) was found between the two tasks. While Ke [11] found that the power of Theta and Alpha in frontal, central and parietal regions show significant difference in task with different loads in the same task, significant difference was also found in the two tasks considering the power of Theta and Alpha both.

Nonlinear methods have been applied to researches of mental workload [13-14] to cover the shortage that linear analysis cannot effectively describe the complexity and regularity of EEG signal for its obvious nonlinearity [12]. Liu [13] found permutation entropies of EEG in frontal and central regions with scales larger than 15 can effectively distinguish different task loads in a verbal N-Back task. However, the applicability of the nonlinear feature of EEG in different task with different information types still remains uncertain.

In conclusion, there are advantages and disadvantages to both linear and nonlinear analysis methods of EEG in the evaluation of mental workload. Spectral power can reflect the activation level of the brain, nonlinear features such as entropies show the complexity of EEG signals. The combination of these two methods may realize cross-task mental workload evaluation in tasks with different information types more effectively. In this paper, linear and nonlinear features of EEG were extracted and analyzed in working memory load tasks with different information types, their effectiveness of accessing mental workload in different tasks were explored, which can be helpful settling the issue of cross-task mental workload evaluation.

II. METHODS

A. Participants

Eighteen right-handed, healthy subjects (9 males and 9 females, Age: 25.6±2.4 years) with normal or corrected vision participated in the study. None of the participants had color blindness or color weakness. All participants gave written informed consent and the study was approved by Beihang University Ethics Committee.

Asterisk indicates corresponding author.

All authors are with Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, No. 37 Xueyuan Road, Beijing, China, 100191 (Correspondence: hjniu@buaa.edu.cn).

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

B. Experiment Design and Data Acquisition

N-Back paradigm was used and verbal or object items were presented in different positions of the checkboard, depicted in Fig. 1 (A). Specifically, each participant needed to perform four tasks, namely Verbal, Object, Spatial (Verbal) and Spatial (Object) N-Back. In the verbal or object task, participants were guided to judge if the letter (or the object) presented now is matched to the one presented previously without considering their positions; and in the two spatial tasks, positions of the items (letter or object) were the only information need to be considered. Three task loads were set by manipulating the "N" from 1 to 3.

Taking verbal 2-Back task as an example, letters were presented on the screen sequentially every 2.5 seconds, each of them was displayed for 0.5 seconds and then disappeared. Participants should determine whether the present letter is the same as the second letter displayed previously. If so, "←" should be pressed and conversely "→". All the participants performed 40 trials in each task at each level of load, 50% of which were set as matching answers. Response time of each trial was recorded during the experiments, accuracy of each task was calculated and NASA-TLX scale for subjective evaluation was asked to be completed after each task.

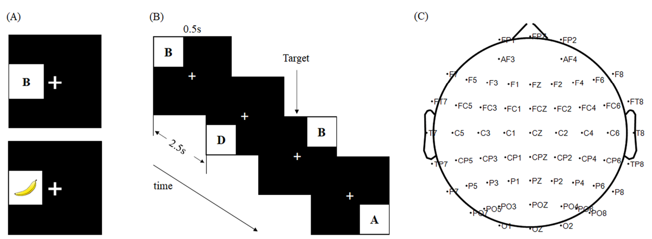

Figure1. N-Back Task and EEG Electrode Locations

The experiments were carried out in a sound-shielded room. EEGs (60 channels, electrode layout was shown as Fig. 1 (C)) were recorded at 1000Hz by using NeuroScan system. The electrode impedances were kept below 10kΩ during recording.

C. EEG Pre-processing

Band pass filter (0.5-40Hz) was applied to eliminate direct component and noises with high frequencies. Eye blinks, vertical/horizontal EOGs and EMGs were removed from the original EEGs by using the auto-recognition algorithm SASICA [15] after resampling signals at 256Hz. Data were segmented into epochs started 500ms before stimulus onset and ended 2000ms after stimulus onset.

D. Feature Extraction

Welch's method was used to calculate spectral power of each epoch at each channel. The relative power of Theta (4-7.5Hz) and Alpha (8-13Hz) were calculated as follows:

$$P_{re}\left(k\right)=S_k \Big/ \sum\nolimits_{j\in\{Delta,Theta,Alpha,Beta\}} S_j \qquad (1)$$

S_k represents power in k spectral, in which k valued as Theta or Alpha. The Theta/Alpha ratio of each channel was also calculated.

Sample entropy can describe the non-linearity of signals, and multiscale entropy analysis [16] can quantify the complexity of time series on different time scales. Multiscale entropy of EEGs in each task was calculated in this paper. A coarse-graining processing was taken on the EEGs of different channels and scales were taken from 1 to 20 to reconstruct new time serials firstly. Then sample entropy at each scale was calculated.

E. Statistical Analysis

Multivariate analysis of variance method was used to analyze the relationship between power in different frequency bands as well as the multiscale sample entropy and mental workload in different tasks with various task loads. All the analysis was carried out by using SPSS 24.0.

III. Results

A. Behavior Results

NASA-TLX scores of different tasks were shown in Fig. 2. As expected, subjective scores in the four tasks increased with the task loads. Two-way ANOVA (4 task types × 3 task loads) was performed and main effect of task load was observed ($p<0.05$). Multiple comparisons show that scores of tasks with different loads were significantly different from each other (1-Back vs. 2-Back: $p<0.05$, 2-Back vs. 3-Back: $p<0.05$) in all the four tasks.

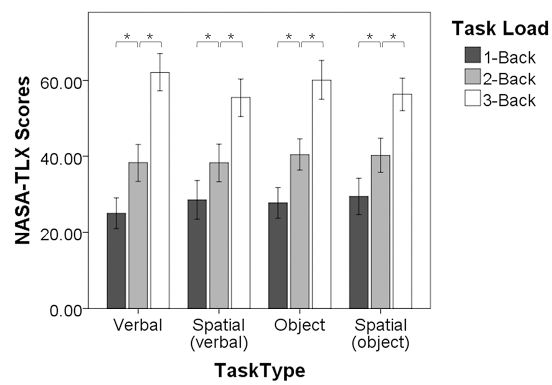

Figure 2. Average Subjective Scores for Each Task Condition for All Participants (Annotations * means $p<0.05$)

Fig. 3 illustrated the mean accuracies and response time in the four tasks. It can be found that with the increment of task loads, the accuracy decreased and the response time became longer. Main effect of task load on both accuracy and response time were observed. Accuracies and response time of tasks with different loads were both significantly different from each other (for accuracy: 1-Back vs. 2-Back: $p<0.05$, 2-Back vs. 3-Back: $p<0.05$; for response time: 1-Back vs. 2-Back: $p<0.05$, 2-Back vs. 3-Back: $p<0.05$). Main effect of task type on response time was also found.

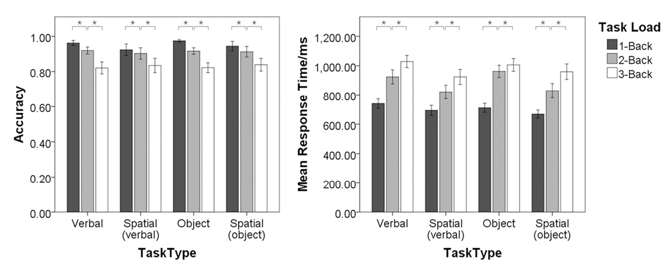

Figure 3. Average Accuracy and Response Time for Each Task Condition for All Participants (Annotations * means $p<0.05$)

B. Spectral Analysis

The brain area and their corresponding spectral features that can reflect mental workload were selected firstly. Specifically, power of Theta in frontal and parietal regions, Alpha in central-parietal regions and Theta/Alpha ratio in frontal-central regions show significant differences ($p<0.05$).

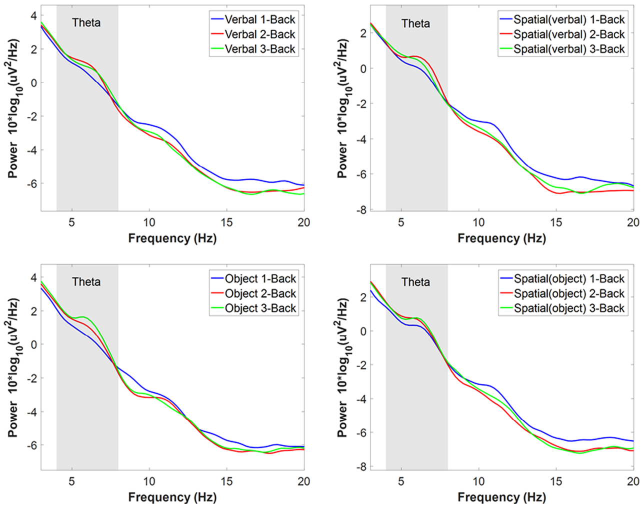

Figure 4. Average EEG spectra over the 3-20Hz interval for each task condition at Fz for all participants

Power spectra in the frontal (Fz) and parietal (Pz) regions were illustrated in Fig. 4 and Fig. 5. Results from ANOVA shows that with the increment of task loads, power of Theta band increased significantly first and decreased slightly then (1-Back vs. 2-Back: $p<0.05$, 2-Back vs. 3-Back: $p>0.05$) in all the four tasks; On the contrary, the power of Alpha band decreased significantly first and increased then (1-Back vs. 2-Back: $p<0.05$, 2-Back vs. 3-Back: $p>0.05$). Fig. 6 displayed the Theta/Alpha ratio in frontal regions (Fz). It can be seen that the Theta/Alpha ratio presented the same regularity as power of Theta band.

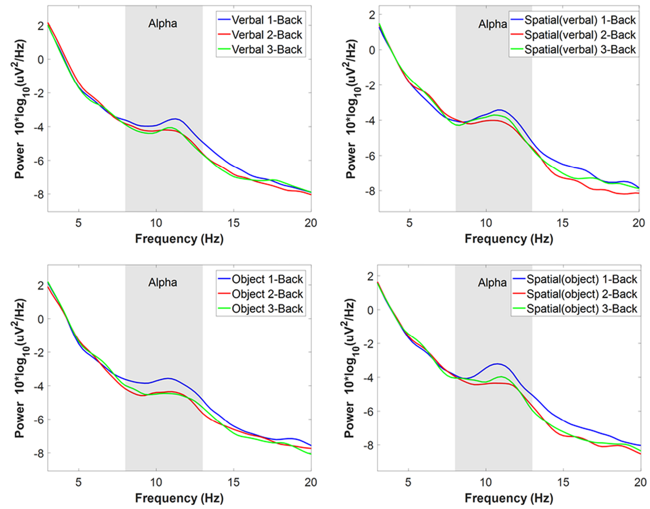

Figure 5. Average EEG Spectra over the 3-20Hz Interval for Each Task Condition at Pz for All Participants

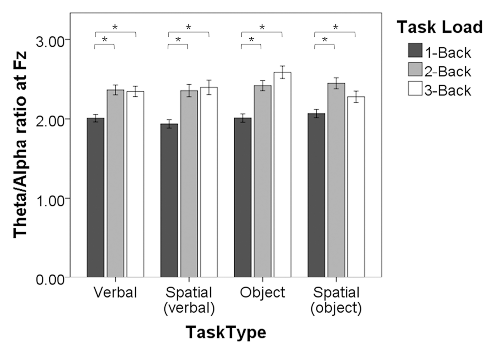

Figure 6. Average Theta/Alpha Ratio for Each Task Condition at Fz for All Participants (Annotations * means $p<0.05$)

TABLE I. MULTISCALE SAMPLE ENTROPY FOR EACH TASK CONDITION AT PZ

		Verbal	Object	Spatial (Verbal)	Spatial (Object)
Scale 1	1-Back	0.2460±0.1302	0.2341±0.1355	0.2673±0.1268	0.2536±0.1303
	2-Back	0.2574±0.1306	0.2744±0.1268*	0.2775±0.1219	0.2860±0.1206*
	3-Back	0.2623±0.1299	0.2630±0.1285*	0.2882±0.1288*	0.2919±0.1236*
Scale 5	1-Back	0.6136±0.3044	0.5846±0.3161	0.6563±0.2859	0.6642±0.2811
	2-Back	0.6285±0.2952	0.6724±0.2898*	0.6886±0.2733*	0.7025±0.2745*
	3-Back	0.6436±0.2925	0.6463±0.2995*	0.7037±0.2819*	0.7197±0.2721*
Scale 10	1-Back	0.6553±0.3160	0.6353±0.3397	0.7029±0.3047	0.7118±0.2925
	2-Back	0.6877±0.3202	0.7325±0.3050*	0.7471±0.2883*	0.7709±0.2950*
	3-Back	0.7018±0.3151*	0.7060±0.3267*	0.7641±0.2973*	0.7958±0.3002*
Scale 15	1-Back	0.7001±0.3363	0.6679±0.3482	0.7459±0.3299	0.7667±0.3234
	2-Back	0.7435±0.3547*	0.7934±0.3408*	0.8074±0.3293*	0.8318±0.3328*
	3-Back	0.7526±0.3414*	0.7652±0.3575*	0.8253±0.3336*	0.8517±0.3265*
Scale 20	1-Back	0.7385±0.3633	0.7205±0.3840	0.7870±0.3612	0.8119±0.3592
	2-Back	0.7858±0.3766*	0.8427±0.3814*	0.8503±0.3593*	0.8875±0.4428*
	3-Back	0.8021±0.3772*	0.8052±0.3840*	0.8703±0.3882*	0.9119±0.4548*

Comparations between different types at each task load were conducted and showed that there are no significant differences for frontal Theta power and central-parietal Alpha power between verbal and object task, as well as between two spatial tasks ($p>0.05$), while the differences between content-specific (verbal and object tasks) and spatial tasks were significant ($p<0.05$). No difference in Theta/Alpha ratio was observed in all the four tasks ($p>0.05$).

C. Multiscale Sample Entropy Analysis

Sample entropies at scale more than 10 of parietal EEG show a main effect on task load in all the four tasks ($p<0.05$). In the object and spatial tasks, meanwhile, significant differences were also found in sample entropies in frontal area (scales more than 10), central area (all the scales) and parietal area (all the scales).

Table 1 shows the sample entropies of parietal (Pz) EEGs at each scale, in which "*" means there's significant difference between 1-Back and 2-/3-Back. Sample entropies on scales more than 10 in verbal 2-Back and 3-Back tasks were found to be larger than that in 1-Back task ($p<0.05$), while difference was not observed between 2-Back and 3-Back tasks ($p>0.05$). In object and the two spatial tasks, sample entropies at almost all scales in 2-Back and 3-Back tasks were significantly larger than that in 1-Back task ($p<0.05$), difference between 2-Back and 3-Back was not observed either.

No differences were found in verbal and object tasks considering the factor of task type. Sample entropies (scales$\geq$14) in spatial (object) task were larger than that in spatial (verbal) task ($p<0.05$), and sample entropies at each scale in verbal and object tasks were smaller than that in spatial tasks ($p<0.05$).

IV. Discussion

By designing N-Back tasks of different information types, regularities of EEG linear feature (spectral power in different frequency bands) and nonlinear feature (multiscale sample entropy) in different tasks with various loads were studied in this paper. The effectiveness of accessing cross-task mental workload with these features was analyzed meanwhile. Results showed that EEG power in frontal Theta, central-parietal Alpha, frontal-central Theta/Alpha ratio and sample entropies at scales more than 10 in parietal regions can reflect mental workload levels effectively. Regularities of power in frontal Theta, central-parietal Alpha and sample entropies at scales more than 10 in parietal regions in verbal task are consistent with those in object task, which indicated that these features are appliable in these two tasks. Theta/Alpha ratio has a consistency between the four tasks and may has the potential to be used in the cross-task mental workload evaluation.

The frontal region of human brain was considered an important component of the attentional system and the Theta rhythm in this region was deemed to be associated with attention and concentration [10].The rise of power in frontal theta observed in this paper indicated an intensified mental effort was required when participants engage in more complex and attention-demanding tasks. The Alpha rhythm is thought to be generated at widespread area of cortex and appears to reflect a relaxed awareness state of human brain. Decline of power in Alpha band may mean the transition of brain state from "idling" to "active" when the task load was higher. Description of complexity by using sample entropy is based on the information measurement of time series . The larger sample entropy observed in tasks with higher load may indicate more complex information processing in the brain. The significant differences between content-based tasks and spatial tasks suggest more attention and mental effort are needed when processing spatial information.

The slight difference between 2-Back and 3-Back tasks shown by ANOVAs may attributed to the constraint of working memory [16] that the response of human brain approached a plateau when task load met or exceeded the capacity. Deeper exploration should be carried out to inspect the general applicability of the linear and nonlinear features in tasks with other information.

References

[1] Heard, Jamison, Caroline E. Harriott, and Julie A. Adams, "A survey of workload assessment algorithms," IEEE Transactions on Human-Machine Systems, vol. 48, no.5, pp. 434-451, 2018.

[2] Paxion, Julie, Edith Galy, and Catherine Berthelon, "Mental workload and driving," Frontiers in psychology, vol. 5, pp. 1344, 2014.

[3] Young, Mark S., et al., "State of science: mental workload in ergonomics," Ergonomics, vol. 58, no. 1, pp. 1-17, 2015.

[4] Astrand, Elaine, "A continuous time-resolved measure decoded from EEG oscillatory activity predicts working memory task performance," Journal of neural engineering, vol. 15, no. 3, pp. 036021, 2018.

[5] Heine, Tobias, et al., "Electrocardiographic features for the measurement of drivers' mental workload," Applied ergonomics, vol. 61, pp. 31-43, 2017.

[6] Foy, Hannah J., and Peter Chapman, "Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation," Applied ergonomics, vol. 73, pp. 90-99, 2018.

[7] Baldwin, Carryl L., and B. N. Penaranda, "Adaptive training using an artificial neural network and EEG metrics for within-and cross-task workload classification," NeuroImage vol. 59, no. 1, pp. 48-56, 2012.

[8] Brouwer, Anne-Marie, et al., "Estimating workload using EEG spectral power and ERPs in the n-back task," Journal of neural engineering, vol. 9, no. 4, pp. 045008, 2012.

[9] Wang, Shouyi, Jacek Gwizdka, and W. Art Chaovalitwongse, "Using wireless EEG signals to assess memory workload in the $ n $-back task," IEEE Transactions on Human-Machine Systems, vol. 46, no. 3, pp. 424-435, 2015.

[10] Gevins, Alan, et al., "High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice," Cerebral cortex, vol. 7, no. 4, pp. 374-385, 1997.

[11] Ke, Yufeng, et al., "Towards an effective cross-task mental workload recognition model using electroencephalography based on feature selection and support vector machine regression," International Journal of Psychophysiology, vol. 98, no. 2, pp. 157-166, 2015.

[12] Stam, Cornelis J, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," Clinical neurophysiology, vol. 116, no. 10, 2266-2301, 2005.

[13] Liu, Po-Kang, et al., "Entropy and Complexity Assisted EEG-based Mental Workload Assessment System," 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS), pp. 1-4, 2019.

[14] Zarjam, Pega, et al., "Estimating cognitive workload using wavelet entropy-based features during an arithmetic task," Computers in biology and medicine, vol. 43, no. 12, pp. 2186-2195, 2013.

[15] Chaumon, Maximilien, Dorothy VM Bishop, and Niko A. Busch, "A practical guide to the selection of independent components of the electroencephalogram for artifact correction," Journal of neuroscience methods, vol. 250, pp. 47-63, 2015.

[16] Zhang, Dan, et al., "Functional connectivity among multi-channel EEGs when working memory load reaches the capacity," Brain research, vol. 1631, pp. 101-112, 2016.